Data Science Fellowship Wrap-up

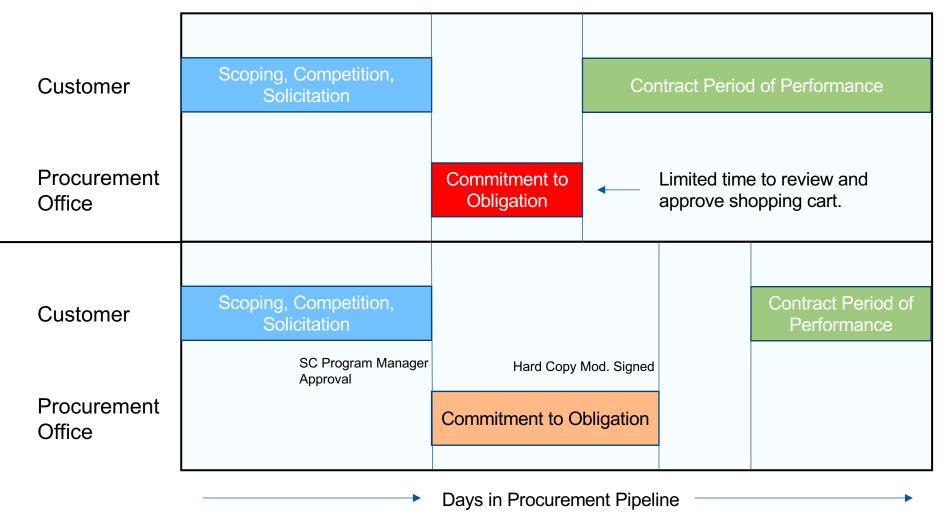
Lessons Learned



Learning About Procurement Through Tableau Work



Close Call Shopping Carts Procurement Pipeline





What is the focus?

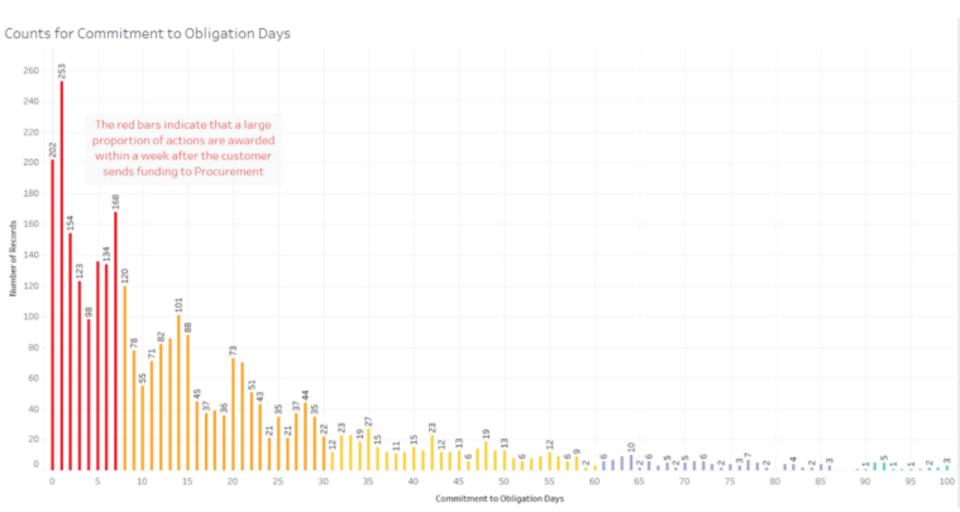
We looked at shopping carts for which the commitment occurs a week or less before the obligation.

What are the goals?

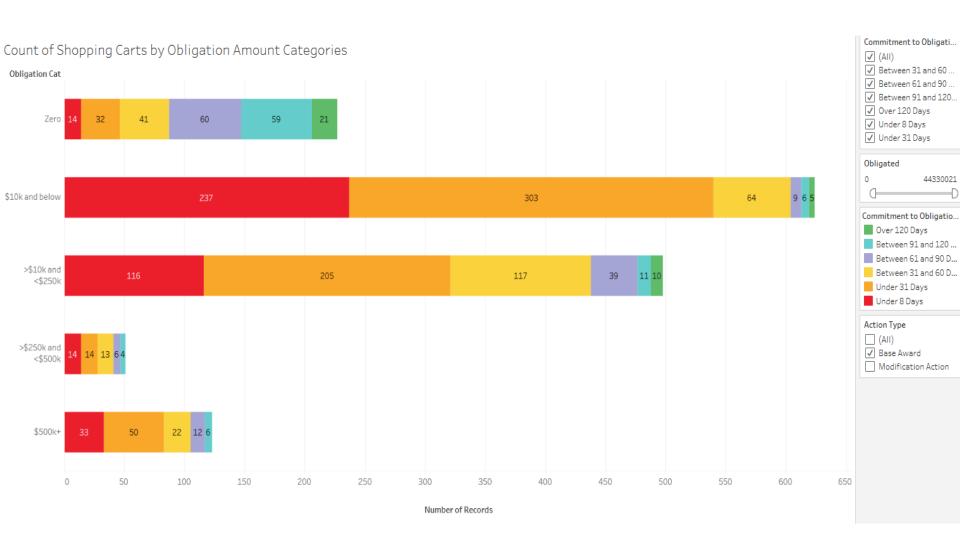
Evaluate whether Program Offices are committing work timely. More specifically, look for trends between the date committed and the date obligated.

Evaluate whether we are increasing our risk by allowing Program Offices to keep the funding under later in the fiscal year.

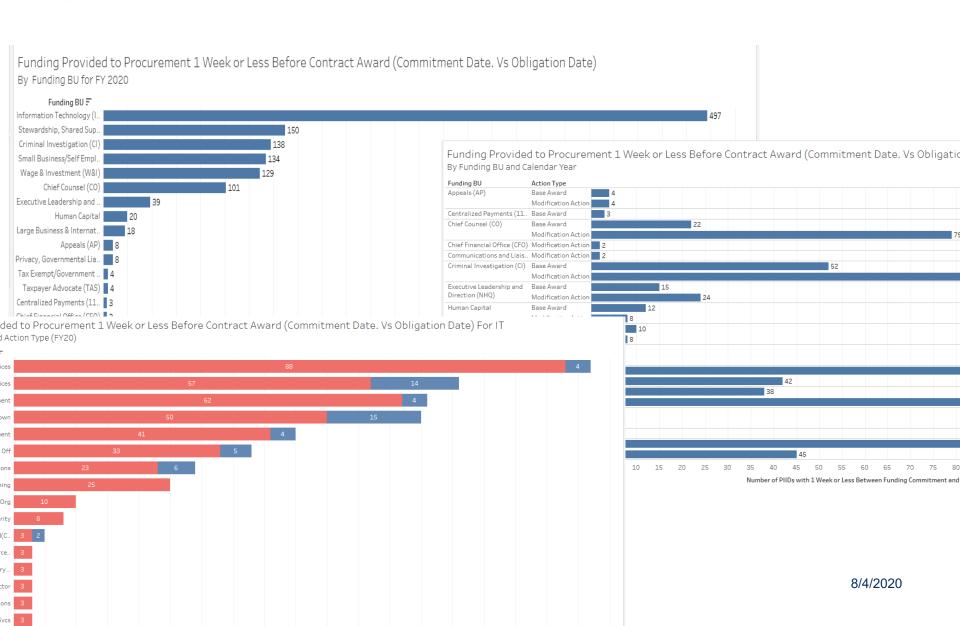














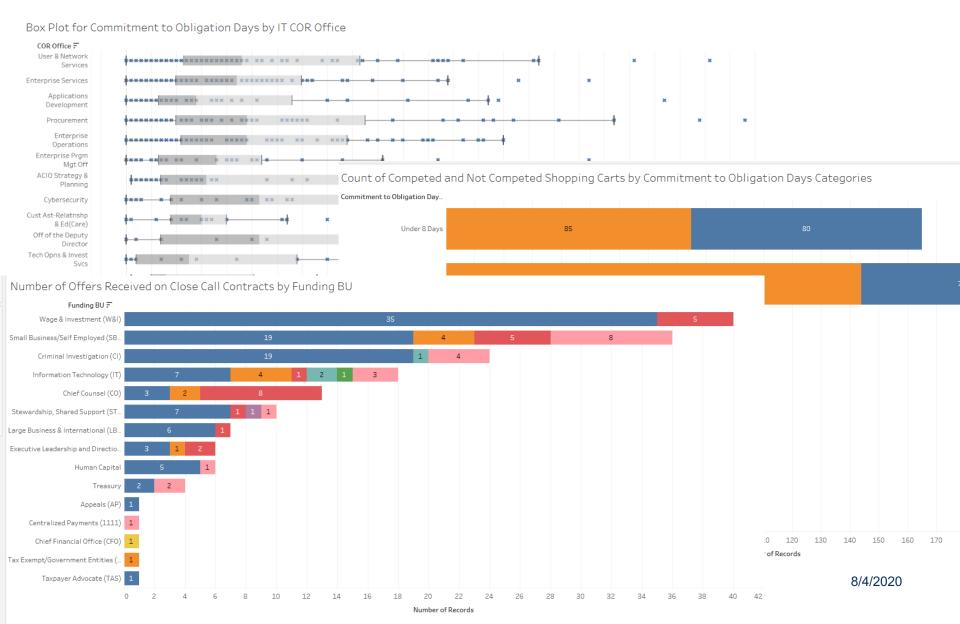




Tableau / Visualizations Takeaways

- Reach out to experts to define motivations of your project
- Start at the top, people's brains don't work backwards
- Two simple graphs and a slow walkthrough are better than one mess and a confused stakeholder
- Documentation and note taking is a constant process
 - Always ask for feedback
- You're not too smart to ask questions



Moving on to Models

Close Call Follow Up Question:

Bridge Contracts. What percentage of close calls resulted in bridge contracts? (Note: we may limited data on bridge contracts. Some COs may label these in the description)



An extension to an existing contract beyond the period of performance (following exercise of all options meeting the requirements of FAR 17.207)

Example:

Procurement Pipeline - Project Scoping - Vendor Sourcing - Competitive Contracting Original Award - Period of Performance - Contract Value - Scope of Work Modification Action - Extended period of performance (usually > 6 months) - Scope of work expanded - Contract value increased

But we still cannot know which modification actions are bridge mods.





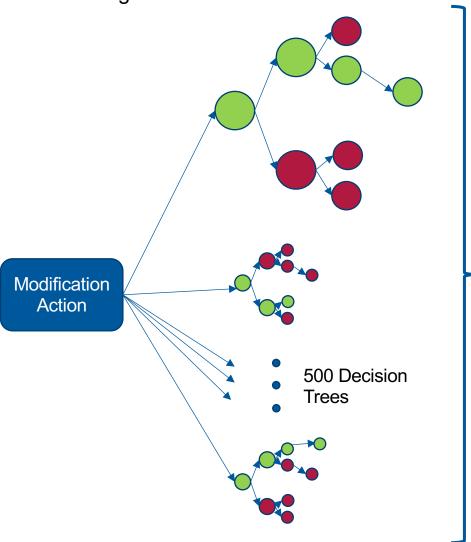
An extension to an existing contract beyond the period of performance (following exercise of all options meeting the requirements of FAR 17.207)

Award ID	Description	End Date	Days Extended	Value	Raised Value?	Total Raise in Value	Contract Length	Is a Bridge?
0001-101	Translation Services	01-30-2020	N/A	\$1000	N/A	\$0	365 Days	N/A
0001-101	Translation Services	01-30-2020	0	\$2500	True	\$1500	365 Days	False
0001-101	Translation Services	05-30-2020	123 Days	\$5000	True	\$2500	488 Days	True
2002-202	Construction Services	02-01-2020	N/A	\$100,000	N/A	\$0	100 Days	N/A
2002-202	Construction Services	08-05-2020	186 Days	\$350,000	True	\$250,000	286 Days	False

- No labelled data
- No easy "catch-all" solution
- Data elements for exercise of options are missing in USASpending



Training a Random Forest Model



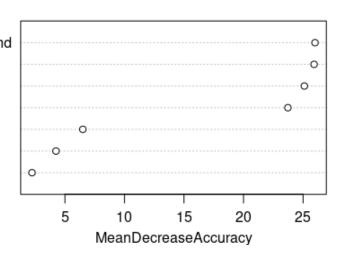
Label Decision Function

Ex. > 250 Trees => Bridge

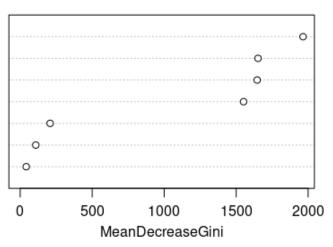


Random Forest Feature Importance

extension_beyond_original_contract_end base_and_all_options_value value_raised extended_beyond_6_months total_raise_in_value contract_length affecting action type



base_and_all_options_value
extension_beyond_original_contract_end
value_raised
extended_beyond_6_months
total_raise_in_value
contract_length
affecting_action_type





Bridge Contracts

New, interim sole-source contract awarded to the same or a new contractor to cover the timeframe between the end of the existing contract and the award of a follow-on contract

Example:

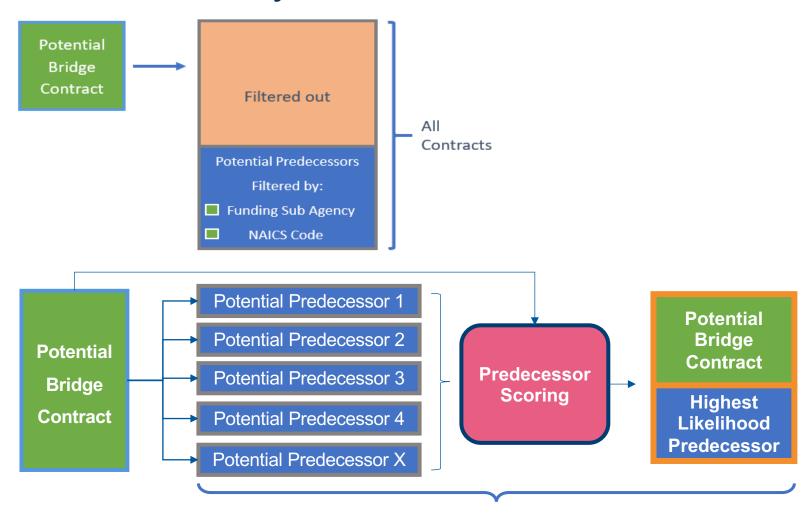
Expiring Contract

- Same or Different Vendor as Expiring Contract
- Immediately Following Expiring Contract
- Limited Sourcing
- Similar Award Description
- Same Funding Agency + NAICS code

Follow On Contract



Bridge Contracts Predecessor Analysis



Parallelized



Predecessor Scoring

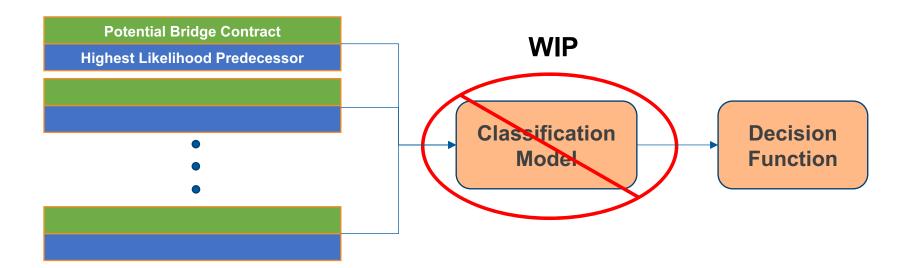
Method	Output		
Contract Description Predecessor Description	Cosine Description Similarity $similarity(doc_1, doc_2) = cos(heta) = rac{doc_1 doc_2}{ doc_1 doc_2 }$		
Start Date End Date	Period of Performance Gap		
Vendor vs. Vendor	Same Vendor Boolean		

Predecessor Score

= Cosine Description Similarity + $\{Same\ Vendor?\ 1:0\} + (1 - \frac{PoP\ Gap}{7})$



Bridge Contracts Modelling



Still need:

- Predecessor Labels
- Bridge Contract Labels



Bridge Modifications and Contracts Takeaways

- Solve sub problems first, build tools for future work
- If you don't have labelled data, that's okay
- Be clear when communicating your project and work
- Documentation should be continuous, don't wait until the end
- Don't hijack Mercury to parallelize just because you're impatient



Questions?